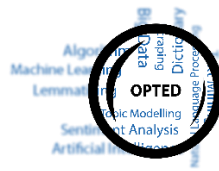




University
of Exeter



**A survey of the research
challenges and training
needs in text analysis
for political science
research**

A survey of the research challenges and training needs in text analysis for political science research

21 November 2022

1. Citation

Scotto di Vettimo, M., Haraldsson, A., Gelovani, S., Banducci, S., Kalsnes, B., Koc-Michalska, K., and Theocharis, Y. (2022). A survey of the research challenges and training needs in text analysis for political science research. [Data collection]. DOI: [10.24378/exe.4344](https://doi.org/10.24378/exe.4344)

2. Description

This online survey assessed the research challenges, training needs and preferences of researchers and practitioners (of all seniority and skills level) that use or are interested in using text analysis techniques in their research. The survey has been conducted within the context of the Horizon 2020 OPTED project (<http://www.opted.eu>), of which all the authors are part. The survey has been approved by the Research Ethics and Governance Office of the University of Exeter (REF: 509745).

2.1. *Population and sample recruitment*

The target population represents the user community that the OPTED project aims to serve. More specifically, we identified as the primary target group academic researchers and scholars that work in the realm of text analysis. However, a second important target group of OPTED are academic researchers with an interest in text-as-data approaches but limited knowledge or resources to invest in such methodologies. Therefore, we wanted to assess the research challenges and training needs of a very heterogeneous research community that comprises PhD students, early- and mid-career researchers as well as senior researchers. Also, we sought to reach researchers that are interested in text analysis, though have not used it already in their research.

To recruit our respondents, we employed a three-pronged approach relying on: (i) lists of contact details gathered by other work packages of the OPTED network for research purposes; (ii) lists of contacts and mailing lists of individual members of the OPTED network; and (iii) mailing lists of organisers of events related to text analysis.

First, other OPTED work packages had previously collected contact details of scholars and practitioners working on text analysis. The specific inclusion criteria varied from one work package to another. For instance, WP6 targeted authors who published any studies using quantitative text-based research over the past five years in top journals in political science, communication, sociology, and psychology, whereas WP2 collected email addresses of authors studying citizen-produced political texts. These populations are contained within our primary target group and represent a key part of the wider user community.

Secondly, to avoid the oversampling of scholars researching on specific types of political texts, we also asked all OPTED members to disseminate the survey to all the list of contacts that they had in their possession. In some cases, this means to reach out to mailing lists

constructed for research purposes other than OPTED, or list of participants to classes, workshops or other events organised by the member of the OPTED network, or simply colleagues potentially interested in text analysis.

Thirdly, to reach a second important target group – researchers that have not already employed text analytical techniques, but that are nevertheless interested in these approaches – we decided to contact organisers of the main events related to text analysis held in Europe from 2018 up to March 2022 (e.g., method schools, conferences, or workshops), and to ask them to share the survey with all the event participants. Additionally, we also asked members of the steering committees of some ECPR Standing Groups to circulate the survey to the group members. Section 3 provides the list of events and ECPR groups contacted. By contacting training events and workshops, we also make sure that we reach more junior researchers like PhD students, who may not be included in other lists created on the basis of academic publications.

2.2. *Dimensions of training needs analysed*

The survey included seven sections (see also questionnaire in Section 4). We focused on different dimensions of training and research needs. The first section contained one question to understand what types of political texts the respondents used in their research. Respondents that do not use and do not even plan to use any textual data are brought to the end of the survey. The second section consists of four questions tapping into the type of text analysis methodology and software used (or planned to use). Based on whether they use computation or non-computational methodologies, respondents are then prompted with more targeted questions about specific tools used and challenges faced when doing text analysis. A third and fourth section (three and two questions, respectively) asks questions about data access and publishing and gaps in database improvements. The fifth section contains two questions about challenges related to multilingual text analysis research. The sixth section consists of fifteen questions about existing and future training opportunities, channels used to acquire new skills and participation in and evaluation of existing training events. Finally, the last section asks five questions related to the demographic characteristics. Although most of these questions had predetermined answer options, from a best practices point of view we also included qualitative questions to allow participants to express needs and preferences that were not covered in the set survey options, so as to capture more nuanced feedback and comments.

2.3. *Dissemination*

The survey was programmed on Qualtrics, and the dissemination of the questionnaire started on February 23, 2022, when we asked OPTED members and event organisers to forward the survey to their lists of contacts. We sent out a reminder 14 days later and we also shared the survey on social media various times.

3. List of text analysis events and ECPR Standing Groups contacted

A. Text analysis events

- COMPTTEXT Conference (2018, 2019, 2020, 2022)
- Computational Text Analysis Workshop – Florence (2020)
- ECPR Summer School – Text analysis courses (2018, 2019, 2020, 2021)
- ECPR Winter School – Text analysis courses (2018, 2019, 2020, 2021)
- Essex Summer School – Quantitative text analysis course (2020, 2021)
- Quantitative Text Analysis Workshop – Dublin (2019)

B. ECPR Standing Groups

- Elites and Political Leadership
- European Union
- Interest Groups
- Internet and Politics
- Laws and Courts
- Parliaments
- Participation and Mobilization
- Political Communication
- Political Methodology
- Political Networks
- Political Parties
- Political Representation
- Politics and Technology
- Regulatory Governance

4. Questionnaire

Dear Colleague,

This survey is part of the European Union funded research project OPTED: “Observatory for Political Texts in European Democracies: A European research infrastructure” ([Horizon 2020 Grant 951832](#)).

This survey is aimed towards social scientists who use text analysis methods in their research. It takes about 15 minutes to complete. You will be asked about your experience with using text analysis methods and, in particular, about your experience with existing resources and materials, and preferences for additional training resources. The purpose of the survey is to learn more about the challenges of conducting text analysis research. The outcomes of this survey will be used to create actionable insights to advance the state of text analysis in the social sciences to the benefit of the social science community.

Before you start the study, it is important that you are aware of the voluntary nature of the survey and how we protect your privacy. All the data is anonymous and will be treated in an aggregated manner. Please read information available at this [link](#) and do not hesitate to ask for clarification if you have questions (see contact details below).

By clicking on "I consent", you confirm the following: I am 18 years or older. I have read and understood the information for participants. I agree to participate in the research and to the use of data obtained from it. I reserve the right to revoke this consent without giving any reason. I reserve the right to discontinue the research at any time.

Thank you for your participation in this study!

For questions please contact: Prof. Susan Banducci at the University of Exeter (s.a.banducci@exeter.ac.uk) or Prof. Karolina Koc-Michalska at the Audencia Business School (kkocmichalska@audencia.com).

consent

If you do not consent to participate in the study, you can now end your participation by clicking on "I prefer not to participate".

- I consent
- I prefer not to participate

End of Block: Informed Consent

Start of Block: Uses text data or not

q1

First, we would like to learn about your use of textual data in your research, as well as the methods that you use to study text.

What sorts of text do you use, plan to use or wish to use in your research? Please consider any uses of textual data in your research, even if this may not be your primary research focus.

| | Don't use and no plans to use | Don't use but would maybe use in future | Currently using or used in the past | Don't know |
|--|-------------------------------|---|-------------------------------------|-----------------------|
| legislative, bureaucratic and government organizations (e.g., parliamentary speeches, bills, laws) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| political organizations such as political parties (e.g., manifestos, speeches, press releases) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| individual politicians (e.g., speeches, interviews, social media posts) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| journalists and/or published by mass media outlets (e.g., news articles, op-eds) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| non-media commercial organizations such as enterprises, lobby groups (e.g., reports, campaign materials, press releases) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| non-governmental organisations and social movements (e.g., forums, pamphlets, campaign materials) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| individual citizens (e.g., tweets, Facebook posts, blog posts, petitions) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Other (please specify): | | | | |

End of Block: Uses text data or not

Start of Block: Methods & Texts

q2

To what extent do you use or wish to use the following text analysis methods in your own research?

| | I don't use it | I would like to use it in the future | I only use it while collaborating with others | I rarely use it | I regularly use it |
|---|-----------------------|--------------------------------------|---|-----------------------|-----------------------|
| Qualitative text analysis (e.g., discourse analysis, conversation analysis) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Quantitative manual text analysis (e.g., manual content analysis) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Computational text analysis (e.g., automated content analysis, machine learning, text mining) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

q2_open

If you use other methods than those mentioned above, please specify:

q4

Please indicate how much you agree or disagree with the following statements, relating to reliability testing (such as, e.g., inter-coder reliability tests).

| | Strongly disagree | Somewhat disagree | Somewhat agree | Strongly agree | Don't know |
|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| It is important to conduct a reliability test | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Conducting reliability tests is difficult | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| It is at times unclear what reliability test should be used | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| I always report reliability in papers based on text data | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Academic journals require reliability tests for the type of text analysis I conduct | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

End of Block: Methods & Texts

Start of Block: IF Computational

comp1

Which kinds of computational tools do you use to study text?

| | I'm not familiar with | I know, but did not apply | I have applied |
|---|-----------------------|---------------------------|-----------------------|
| Text statistics (e.g., word frequencies) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Automated extraction (e.g., link/hashtag extraction, concordances, keywords-in-context) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Dictionaries/keyword searches | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Natural language processing tools (e.g., Part-of-Speech taggers, dependency parsers) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Sentiment scoring | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Semantic network tools | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Topic models/text clustering tools | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Word embeddings | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Text similarity scoring (e.g., plagiarism software) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Document scoring (e.g., Wordfish, Wordscore, Wordshoal) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Supervised machine learning | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Machine translation | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

comp1_open

I use other computational tools (please specify):

comp3

In your experience, what challenges did you encounter (or you think you might encounter in the future) when using computational methods for text analysis?

| | not a challenge | a minor challenge | a major challenge |
|--|-----------------------|-----------------------|-----------------------|
| Time/effort required (e.g., technical requirements, experience) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Funding required (e.g., for training, fees, licenses) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Availability of required training | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Training is available but you have no capacity to engage | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Limited methodological guidance/documentation of tools | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Infrastructure is not available (e.g., archives, computers with access to specialist software, etc.) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Availability of suitable computational tools for specific measurement purposes | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Availability or comparability of suitable computational tools for the language(s) that I study | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Issues concerning measurement validity/limited nuance | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Skepticism (of myself or others) toward computational methods | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

comp3_open

Other challenges (please specify):

comp4

To what extent was the required time and effort challenging because of the following? (Please tick all that apply)

- Childcare/other caring commitments
- Other personal commitments
- High commitment to research work
- High teaching load
- Other professional commitments
- Other, please specify:
- None of the above

comp6

Please indicate your agreement with the following statements:

| | Strongly disagree | Somewhat disagree | Neither agree nor disagree | Somewhat agree | Strongly agree |
|--|-----------------------|-----------------------|----------------------------|-----------------------|-----------------------|
| I can easily discuss problems related to computational text analysis with colleagues from my department | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| I easily solve most problems related to resources for computational text analysis simply by consulting the documentation made available with the relevant resource | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| I rely on public platforms (e.g., Stack Overflow) to discuss problems related to computational text analysis methods and resources | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Sometimes it is hard to find a user community with which I can discuss problems related to computational methods | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| It is necessary to have computational text analysis skills in order to be competitive on the academic job market | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

comp8

How important would the following be for your research activities?

| | Not at all important | Slightly important | Moderately important | Very important | Extremely important |
|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| A single platform for text analysis tools and resources | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Access to text analysis tools for programming | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

End of Block: IF Computational

Start of Block: IF Not Computational

nocomp1

What challenges were relevant (or might be relevant in the future) for your choice to not use computational methods for text analysis?

| | not a challenge | a minor challenge | a major challenge |
|--|-----------------------|-----------------------|-----------------------|
| Time/effort required (e.g., technical requirements, experience) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Funding required (e.g., for training, fees, licenses) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Availability of required training | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Training is available but I have no capacity to engage | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Limited methodological guidance/documentation of tools | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Infrastructure is not available (e.g., archives, computers with access to specialist software, etc.) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Availability of suitable computational tools for specific measurement purposes | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Availability or comparability of suitable computational tools for the language(s) that I study | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Issues concerning measurement validity/limited nuance | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Reviewers'/editors' skepticism toward computational methods | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Skepticism (of myself or others) toward computational methods | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

nocomp1_open

Other challenges (please specify):

Display This Question: If nocomp1 = 1 [2] Or nocomp1 = 1 [3]

nocomp2

To what extent was the required time and effort challenging because of the following? (please tick all that apply)

- Childcare/other caring commitments
- Other personal commitments
- High commitment to research work
- High teaching load
- Other professional commitments
- Other, please specify:
- None of the above

End of Block: IF Not Computational

Start of Block: Access & Publishing

q5

In your experience, which were challenges that you encountered relating to **accessing text data**?

| | Not applicable | Not a challenge | A minor challenge | A major challenge |
|--|-----------------------|-----------------------|-----------------------|-----------------------|
| Restricted access to data by companies owning or storing the data (e.g., Twitter, Facebook, CrowdTangle, Crimson Hexagon/Brandwatch) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Restricted access to data due to content removal (e.g., online content that is flagged and removed by moderators) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Difficulty finding or identifying all relevant text data (e.g., because words that are spelled correctly by users are not collected) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Difficulty identifying or using the tools needed to access data | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| National ethical research rules or guidelines do not enable accessing certain text | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

q5_open

Other challenges (please specify):

q6

Please indicate how much you agree with the following statements, relating to data privacy, confidentiality, and consent.

| | Strongly disagree | Somewhat disagree | Somewhat agree | Strongly agree |
|---|-----------------------|-----------------------|-----------------------|-----------------------|
| I am aware of the best practices for ensuring confidentiality of those whose text data I collect/analyse | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| I have received sufficient training in the legal and ethical guidelines for data protection and privacy of those whose data I collect/analyse | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| It has not always been possible in my research to ask for consent to participate, even if gaining consent may have been best practice. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

End of Block: Access & Publishing

Start of Block: Databases

q16

In order to undertake political text analysis, researchers need access to textual data (e.g., national legislation, news stories). Thinking about the future, how important are the following improvements to these types of databases for researchers?

| | Not at all important | Slightly important | Moderately important | Very important | Extremely important |
|--|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Expanding the geographical coverage of existing datasets | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Expanding the temporal coverage of existing datasets | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Develop new datasets | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Facilitating the linkage of existing datasets | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Develop public web portals | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Develop procedures for local installations of datasets | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Develop API-web services | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

q16_open

Please provide more details as to what sort of database development is a priority, or suggest other types of database developments to prioritise: _____

End of Block: Databases

Start of Block: Languages

q7

In which language(s) are those texts that you study (or wish to study)? (please tick all that apply)

[Follows boxes with languages]

End of Block: Languages

Start of Block: IF Multiple Languages

lang1

To what extent were the following statements reasons for you to conduct textual analysis in multiple languages? These are languages...

| | Not at all a reason | Minor reason | Major reason |
|---|------------------------|-----------------------|-----------------------|
| that I speak well myself | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| that are spoken in a single country that I study (e.g., Switzerland, India) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| that are particularly relevant for my research | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| in which it is possible/easy to access textual material | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| for which it is possible/easy to find qualified collaborators/ assistants | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| for which it is possible/easy to find suitable tools | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| for which the quality of machine translation suffices | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| for which it is possible/easy to conduct a comparative analysis | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

End of Block: IF Multiple Languages

Start of Block: Training

q8

How did you acquire the skills necessary to utilise text analysis methodologies in your research? (Tick all that apply):

- Self-led online material
- From examples and posts found online
- Free community-led online workshops
- Conference workshops and tutorials
- Peers and colleagues
- Undergraduate/master course
- Postgraduate training as part of my PhD
- Third-party training course
- Other (please specify): _____

q9

Have you ever attended any event (e.g., summer school, workshop, seminars, etc.) related to text analysis?

- Yes
- No

q10

How likely are you to attend a training event related to text analysis in the next 2 years?

- Not likely at all
- Somewhat likely
- Very likely
- Don't know

q11

What is the type of event you would most like to attend in the next two years?

- Offline/in-person
- Online with mostly synchronous sessions
- Online with pre-recorded training materials (e.g., available to watch on demand)
- Hybrid (both online and in-person)
- Other (please specify): _____

q12

If you were to attend a text analysis training session, please tell us whom you would like the most to deliver the training

- Academics from my field who are also using text analysis techniques
- Academics expert in text analysis techniques, but who do not necessarily operate in my field
- Non-academic experts and practitioners (e.g., coming from the industry)

q13

Are you currently involved in the training or supervision of social scientists? (e.g., as part of a degree program, doctoral/postdoctoral supervision, specialized training)

- Yes
- No

Display This Question: If q13 = 1

q13a

Do you believe that, in the context of your training activities for students and social scientists, there are important training needs in relation to text analysis in any of the following areas?

| | Not at all important | Slightly important | Moderately important | Very important | Extremely important |
|---------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Data and open access tools | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Programming and software skills | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Theory and concepts | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Research integrity, ethics | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Display This Question: If q13 = 1

q13a_open

Other training needs (please specify):

q14

As a supervisor of researchers who need to learn text analysis, who do you think is suitable to lead the training sessions for these researchers? (Skip if it does not apply to you)

- Academics from my field who are also using text analysis techniques
- Academics expert in text analysis techniques, but who do not necessarily operate in my field
- Non-academic experts and practitioners (e.g., coming from the industry)

q15

Would you be interested in receiving (further) training in computational text analysis methods?

- Yes**
- No, my skills are sufficient**
- No, I do not plan to use computational text analysis anytime soon**
- No, for other reasons (please specify):**

Display This Question: If q15 = 1

q15a

On what level would you need to receive (further) training?

- On an introductory level
- On an intermediate level
- On an advanced level

End of Block: Training

Start of Block: IF Training events

train1

What kind of events did you attend? (Select all that apply).

- Method school exclusively focused on text analysis methods
- Method school where some courses covered text analysis methods
- Conference exclusively focused on text analysis methods
- Conference where some sessions/panels were dedicated to text analysis methods
- Seminar or workshop organised by another university
- Seminar or workshop organised at my university
- Training event organised by another university
- Training event organised by my university

train2

Overall, how would you rate the events you attended with regard to the following:

| | Not helpful at all | Somewhat unhelpful | Neither helpful nor unhelpful | Somewhat helpful | Very helpful |
|--|-----------------------|-----------------------|-------------------------------|-----------------------|-----------------------|
| Learning new software, resources and techniques | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Improving your training in software, resources and techniques you already know | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Getting to know a community of scholars working with similar methodologies | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Receiving feedback on your ongoing research | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

train3

Thinking about the training events you participated in, what do you think was the key strength which should definitely be kept so as to ensure that they benefit users like you?

train4

Thinking about the training events you participated in, what do you think was the key issue which should be addressed so as to make them more helpful for users like you?

End of Block: IF Training events

Start of Block: OPTED proto-type

q17

The OPTED network is currently working on the proto-type of a platform for text analysis resources for the academic and the non-academic community interested in political texts (you can find more information at:

opted.eu). We highly value any user input on this proto-type, as it will allow us to better meet the needs of the user community we aim to serve.

Which aspects do you think deserve to be prioritised to best satisfy the users' community needs, including your own?

| | Not a priority at all | Definitely a priority | Definitely a priority and I would make use of it |
|--|-----------------------|-----------------------|---|
| An open repository for different types of data sources and relevant documentation | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| An open repository for different types of training materials | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| An open repository for different types of tools, software and packages | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| A platform where users can find recommendations for relevant resources for their research | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| A platform users can contribute to with their own resources (e.g., new databases, software or packages) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| A platform with tools and standards for the validation of computational methods | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| A platform that can be used to work collaboratively on the discovery, creation and sharing of text analysis resources (e.g., codes or data) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| A platform that can host discussions about text analysis resources | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| A platform for where users can access training materials (e.g., videos, slides, Shiny apps) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| A platform that can be used to “re-appraise” existing resources , highlight and solve issues (e.g., bugs) and track updates | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

q18

What do you think are the key features the platform should possess in order to satisfy the user community needs?

End of Block: OPTED proto-type

Start of Block: Demographics

field

What is your main area of research or study? (Please tick all that apply)

- Communications
- Economics
- Political Science
- Psychology
- Sociology
- Other:

country

In what country is the university or employer with which you are affiliated?

▼ Afghanistan (1) ... Zimbabwe (1357)

rank

I am...

- a PhD student
- an early-career researcher (<5 years since PhD)
- a mid-career researcher (5-15 years since PhD)
- a senior researcher (>15 years since PhD)
- Other (please specify):

gender

I identify as

- a man
- a woman
- I do not identify as a man nor as a woman
- I prefer not to say

mailing_list

Would you like to be contacted in the upcoming weeks with more information about the OPTED platform, so as to remain updated on its development and to have the possibility to try and comment on its proto-type?

- Yes, I would like to be added to the OPTED mailing list
- No, I do not want to be added to the OPTED mailing list

End of Block: Demographics